

中国人口增长预测模型

摘要

本文针对中国人口增长的预测问题，建立了改进的矩阵人口增长模型、城乡转移矩阵模型、长期增长预测模型三个模型。同时建立了对人口增长模型进行综合评价的综合评价模型。

在建立模型前我们首先补充了题目并未给出的每年人口总数，处理了有异常的 2003 年生育率等异常数据，完成初步的数据处理工作。

模型一，建立了改进的矩阵人口增长模型。以Leslie提出的分组模型为基础，将男性数量预测同时建立在模型中。在改进模型中死亡率等相关数据使用BP神经网络、灰色预测等方法预测得到不再是固定的。通过模型求解我们得到各指标预测结果。通过对模型的结果分析可知，该离散形式的分组人口增长预测模型具有：对男性数据同步预测中短期更准确、充分利用数据、离散形式便于计算机作数值计算等优点，但模型由于并未考虑城乡人口转移对人口增长的影响，所以只能作为比较优化的粗略中短期模型。

模型二，建立了城乡转移矩阵模型。在模型一的基础上，对人口城乡流动转移产生的转移人口的规模和结构引入了转移矩阵对其进行预测，从而准确的描述了人口转移对整体人口数量、性别比等的影响。在对其中的年龄模式和转移率的预测过程中，我们对最初的线性拟合模型进行了统计检验，发现拟合程度很底。最后确定使用灰色预测和神经网络的预测方法。通过模型求解我们得到表 1 等指标预测结果。对结果进行分析可知，模型二能够考虑到人口流动对人口数量的影响并能够将其有效的量化分析做出转移矩阵，这样就有着更优的中短期预测效果。但模型使用的数据自身具有局限性，所以对于 50 年以上的长期预测，模型的预测仍然不够准确。

表 1 模型二 未来 15 年的人口总数预测结果

单位：十亿人

年份	人口数量	年份	人口数量	年份	人口数量	年份	人口数量
2006	1.3112	2010	1.3440	2014	1.3873	2018	1.4393
2007	1.3177	2011	1.3546	2015	1.3982	2019	1.4588
2008	1.3253	2012	1.3654	2016	1.4104	2020	1.4844
2009	1.3340	2013	1.3762	2017	1.4239		

模型三，建立了长期人口增长模型。根据模糊预测理论，对初始数据进行模糊处理。由于在长期预测时女性数量的稳定性更好，所以以女性的数量变化作为研究对象。引入 Lee-Carter 预测模型对人口的死亡率进行预测；使用 S 型曲线 $u(t) = \frac{1}{1 + e^{\alpha + \beta t}}$ 模型对城镇化进行预测；推导出生育率计算方程；推导出性别比比例的预测模型。在上述数据的预测基础上建立长期增长模型进行预测得到了较为科学准确合理的结论（见表 5 等）。

模型四，建立了综合评价模型。基于层次分析法（见图 16）和模糊评价对人口增长模型预测得到老龄化趋势、出生性别比、城镇化水平等指标进行了综合评价，进而得到何种模型更优的评价结论并通过不同模型的假设可以对政策制定提出一些建议。

最后，我们建立了对产品上架情况进行预测得到人口增长预测结果的扩展模型。

关键词：Leslie 矩阵 人口转移矩阵 模糊评价 层次分析 神经网络

1、问题分析

中国的人口增长问题是一个复杂的问题，对该问题的求解应该考虑循序渐进逐步完善模型。在开始对中国人口增长进行预测建模前我们先要得到每年的人口总数，且处理异常数据。实际上老龄化、出生人口性别比，以及乡村人口城镇化等因素，都影响着中国人口的增长。对中国人口增长的中短期和长期预测，除对人口数量的预测分析外对这些指标也应该进行一定的预测分析。人口增长模型较为复杂性且数据的数量较小(5组)，对其数据的预测运用曲线拟合的效果很可能并不是很好。应该对拟合结果进行统计检验，通过判定系数 R^2 检验、 $D-W$ 检验等统计检验的检验结果来确定是否可以使用曲线拟合。如果拟合度过低，我们就要考虑通过神经网络、灰色预测等方法对数据进行预测。这些方法对数据量要求不是很大且善于处理复杂问题的求解。所以可能更适合对数据的预测。对人口增长做出中短期(15年左右)和长期预测(50年以上)，虽然我们希望能够通过一个模型得到完整的预测，但由于各种预测模型自身的局限使得我们不可能通过一个模型同时得到较为准确的长期预测结果和中短期预测结果，所以特别建立长期预测模型对问题进行求解。人口增长问题的模型可能很多，那么如何对模型进行评价也是我们对此问题解决的一部分。

2、数据处理

在题目所给的数据并未给出每年的人口总数，且所给的数据只是人口抽样数据，而且不一定是1%人口抽样数据，无法推出人口总数。而我们在模型的建立和求解过程中是需要总人数的。所以我们补充了每年人口总数的数据(见附录表7 1987-2005人口数及结构)。同时经观察和相关资料查找可知2003年数据中的生育率部分是有异常的将异常数据进行异常点处理。将异常数据与其它数据进行类比(见图1)发现将异常数据扩大十倍后基本满足数据的大小范围，且满足数据的变化趋势。所以对异常数据进行扩大十倍的处理。

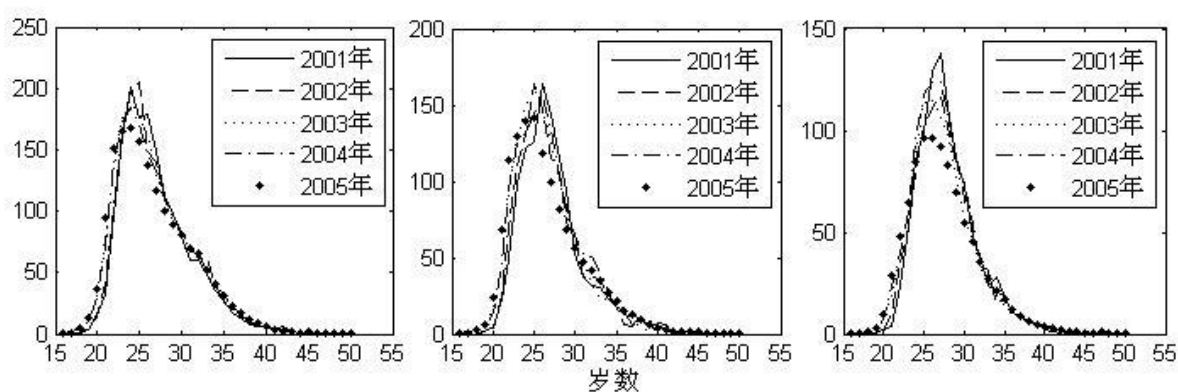


图1 2005-2001 镇-乡-城 处理异常点后的妇女生育率 类比图

3、模型假设

(1) 由于出入境的人口差对全国人口数量的影响是微小的,不妨对于出入境的人口不予考虑。这对于人口基数很大的国家(例如中国)是合理的。

(2) 假设在预测期间不会发生对人口有巨大影响的特殊事件(如大规模战争、严重的瘟疫、大范围自然灾害等)。

(3) 模型均是建立在我国人口政策长期稳定的前提下的。

4、基本符号说明

将城、乡、镇不同性别的人进行编号如下表

表 2 i 编号

类型	城市男性	城市女性	镇男性	镇女性	乡村男性	乡村女性
编号 i	1	2	3	4	5	6

将城、乡、镇三种不同城乡情况分别用编号表示如下表

表 3 j 编号

类型	城市	镇	乡
编号 j	1	2	3

t :代表年份;

k :代表年龄;

$p_j(t)$: t 年, j 类人口的出生性别比;

$b_{jk}(t)$: t 年, k 岁, j 类女性的生育率;

$d_{ik}(t)$: t 年, k 岁, i 类人的死亡率;

$s_{ik}(t)$: t 年, k 岁, i 类人的存活率;

$h_{jk}(t)$: t 年, k 岁, j 类女性的生育模式;

$\beta_j(t)$: t 年, j 类女性的总和生育率;

$x_{ik}(t)$: t 年, k 岁, i 类人的人口总数;

$w_f^s(k, t)$: t 年, k 岁人口分性别的 f 方向的人口转移比;

$X_i(t)$: t 年, i 类人的总人数矩阵 由 $x_{ik}(t)$ 构成一维向量矩阵;

$A_i(t)$: t 年, i 类人的存活矩阵 由 $s_{ik}(t)$ 构成;

$B_j(t)$: t 年, j 类女性的生育模式矩阵;

$W_f^s(t)$: t 年, k 岁, i 类人口城乡转移矩阵;

在本文中用 90 表示 90^+ 。

5、模型的建立和求解

5.1 模型一(改进的矩阵人口增长模型)

5.1.1 基本思路

Leslie 人口模型考虑种群是通过雌性个体的繁殖以雌性个体数量变化为研究对象, 且假设存活矩阵和生育模式矩阵(在本文中将两个矩阵统称为人口变化矩阵)以及男女性别比是固定不变的矩阵人口模型。我们所具有的数据是可以用与雌性相同的理论对男性的数量进行研究的。处于呈城、乡、镇的男性, 从现实角度考虑使其出现男孩的产生应该由相应地区类型的女性繁殖产生。从实际情况我们可以知道由于政策调整作用, 人们生育愿望的改变等因素转移矩阵和男女比例是在随年份变化而改变的。所以我们对矩阵进行预测, 得到随年份变化的人口变化矩阵, 然后利用两个矩阵建立差分方程得到总体的人口增长模型。这样在中短期预测中是更加准确的。同时对男性数量进行预测避免了对男女比例的预测(所给的数据量相对较小进行预测误差较大)要求同时贴近现实、数据利用充分。

5.1.2 模型的建立

(1) 存活矩阵的建立。存活矩阵 $A_i(t)$ 如下:

$$A_i(t) = \begin{bmatrix} 0 & 0 & \cdots & \cdots & 0 \\ s_{i0}(t) & 0 & \cdots & \cdots & 0 \\ 0 & s_{i2}(t) & \cdots & \cdots & 0 \\ \vdots & \vdots & \ddots & & \vdots \\ 0 & 0 & \cdots & s_{i90}(t) & 0 \end{bmatrix}$$

其中, $A_i(t)$ 表示 t 年 i 类人的存活矩阵; $d_{ik}(t)$ 表示在 t 年 k 岁 i 类人的死亡率;

$s_{ik}(t) = 1 - d_{ik}(t)$ 表示在 t 年 k 岁 i 类人的存活率。

(2) 生育模式矩阵的建立。生育模式矩阵 $B_j(t)$ 如下：

$$B_j(t) = \begin{bmatrix} 0 & \cdots & 0 & h_{j15}(t) & \cdots & h_{j49}(t) & 0 & \cdots & 0 \\ 0 & \cdots & \cdots & & \cdots & \cdots & \cdots & \cdots & 0 \\ \vdots & & & & & & & & \vdots \\ 0 & \cdots & \cdots & & \cdots & \cdots & \cdots & \cdots & 0 \end{bmatrix}$$

其中， $B_j(t)$ 表示 t 年 j 类人的生育模式矩阵；女性育龄范围为 $[15, 49]$ ； $h_{jk}(t)$ 由式 $\beta_j(t)h_{jk}(t) = b_{jk}(t)$ 导出。

$b_{jk}(t)$ 表示 t 年 k 岁 j 类女性的生育率； $\beta_j(t)$ 表示 t 年 j 类女性的总和生育率，

$$\beta_j(t) = \sum_{k=15}^{49} b_{jk}(t); \quad \sum_{k=15}^{49} h_{jk}(t) = 1 \quad .$$

(3) 人口总数向量的建立。人口总数向量 $X_i(t)$ 如下：

$$X_i(t) = [x_{i0}(t) \quad x_{i1}(t) \quad \cdots \quad x_{i90}(t) \quad x_{i100}(t)]$$

式中， $X_i(t)$ 表示 t 年 i 类人的总人数矩阵； $x_{ik}(t)$ 表示 t 年 k 岁 i 类人的人口总数， $x_{ik}(t)$ 可由全国总人数的数据和各类人口的比例求出。

(4) 人口增长模型的建立。 $t+1$ 年 i 类人的增加人数是 t 年各年龄繁殖数量之和。由于所有人口的增长都是由女性的繁殖所产生的，男性的数量增加数应由与他处于相同城乡类型的女性的生育转移矩阵、女性前年总数、出生性别比等来确定。根据得到的数据建立模型如下：

$$\begin{cases} X_i(t+1) = A_i(t)X_i(t) + \beta_{\frac{i}{2}}(t)B_{\frac{i}{2}}(t)X_{\frac{i}{2}}(t) \frac{1}{p_{\frac{i}{2}}(t)+1}, & i = (2,4,6) \cdots \cdots (1) \\ X_i(t+1) = A_i(t)X_i(t) + \beta_{\frac{i+1}{2}}(t)B_{\frac{i+1}{2}}(t)X_{\frac{i+1}{2}}(t) \frac{p_{\frac{i+1}{2}}(t)}{p_{\frac{i+1}{2}}(t)+1}, & i = (1,3,5) \cdots \cdots (2) \end{cases}$$

式中， $p_j(t)$ 表示 t 年 j 类人口的出生性别比，在式子中 j 由 i 计算导出。

式(1)是对于 $t+1$ 年女性数量的预测模型，该模型基本与原始的 Leslie 人口模型基本相同，不同的只是在这个模型中原始的 $\beta_{\frac{i}{2}}(t)B_{\frac{i}{2}}(t)X_{\frac{i}{2}}(t)$ 是包括男孩和女孩的孩子数量

增加为了保证求得女儿的数量，要乘上比例系数 $\frac{1}{p_{\frac{i+1}{2}}(t)+1}$ ；

式(2)是对于 $t+1$ 年男性数量的预测模型，其中 $\beta_{\frac{i+1}{2}}$ 、 $B_{\frac{i+1}{2}}$ 、 $p_{\frac{i+1}{2}}$ 等代表了与 i 类男

性相同的城乡类型的女性该年的总和生育率、生育模型矩阵、出生性别比。 $\frac{p_{\frac{i+1}{2}}(t)}{p_{\frac{i+1}{2}}(t)+1}$ 代

表了 t 年该城乡类型女性繁殖生育的所有孩子中男孩的比例。 $A_i(t)$ 代表了 t 年该类型男

性的存活矩阵。 $X_{i+1}(t)$ 代表了 t 年同类型女性的总数矩阵。加数 $\beta_{\frac{i+1}{2}}(t)B_{\frac{i+1}{2}}(t)X_{i+1}(t)\frac{p_{\frac{i+1}{2}}(t)}{p_{\frac{i+1}{2}}(t)+1}$ 整

体表示了由各年龄的女性生殖产生的男性数量的增长。

5.1.3 模型的求解

(1) 存活矩阵的预测。存活矩阵的预测从根本上来讲就是对死亡率 $d_{ik}(t)$ 的预测。

对于死亡率，我们的数据从年份的角度来看实际上只有 5 组，即 2001–2005 的数据，对如此少的数据进行观察并不能对其规律得到很好的结论。如果采用线性曲线拟合误差显然会很大。于是我们采用神经网络的方法对死亡率进行预测。这种方法是根据实际的变化趋势进行预测的，其结果更加准确。

神经网络对复杂非线性系统有曲线拟合能力。基于 BP 神经网络的时间序列预测方法只要输入历史数据进行训练，而自动形成模型。训练误差指标设为 10–12，学习速率设为 0.001，最大步长设为 2000。隐含层中采用双曲正切 S 型传递函数。输出层采用线性传递函数 purelin。使用 matlab 计算可得预测的死亡率并最终得到存活矩阵。

(2) 生育模式矩阵的预测。生育模式矩阵的预测即对出生性别比和生育率的预测。

我们首先对出生性别比进行了线性拟合并对其结果进行了统计检验得到的结论是拟合效果非常不好，于是我们采用带有残差修正的灰色预测法进行预测得到了较为准确的数据。

(3) 人口增长模型的求解。在求解的过程由于模型要求可知 $X_i(t)$ 中 $x_{ik}(t)$ ， k 从 0–91 而实际上数据应该是 0–90，对于第 91 项其存在意义在于保证 90 的数据不被替代，而它的数据可以看作是 100^+ (标记作用) 的数据。其出使值应该为 0，在计算过程中一次求解后将 100^+ 的数据加到 90 的数据上，并将其再次初始化 0，然后进行接下来的递推运算。这样处理从逻辑上是合理的，从具体物理意义上也是正确的。从而得到正确的处理结果。

在得到预测转移随年份变化的预测后我们计算差分方程模型，通过递推得到各类型人口各年龄的预测结果。从而得到整个模型的解答见表 4 模型一 未来 15 年的人口总数预测结果,图 2 模型一 未来 15 年的人口总数预测结果柱形图,图 3 模型一 未来 15 年的人口性别比预测结果柱形图。

表 4 模型一 未来 15 年的人口总数预测结果

单位：十亿人

年份	人口数量	年份	人口数量	年份	人口数量	年份	人口数量
2006	1.3112	2010	1.3431	2014	1.3879	2018	1.4524
2007	1.3179	2011	1.3530	2015	1.4009	2019	1.4762
2008	1.3254	2012	1.3637	2016	1.4159	2020	1.5065
2009	1.3338	2013	1.3753	2017	1.4329		

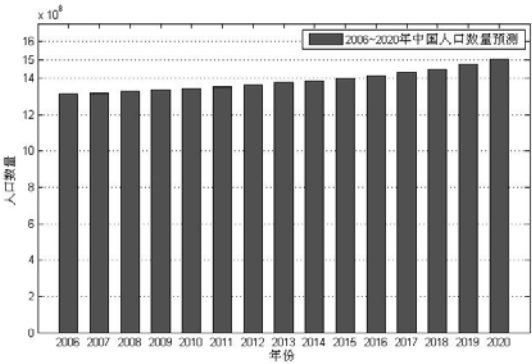


图 2 模型一 未来 15 年的人口总数预测结果柱形图

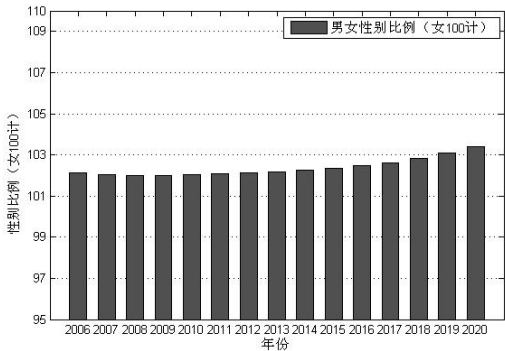


图 3 模型一 未来 15 年的人口性别比预测结果柱形图

5.1.4 结果分析

在本文中对结果的分析，均以附件《国家人口发展战略研究报告》中的相应预测结果为比较准确的预测结果(简称为准确结果)，将所预测的结果与之比对进行评价。以下不再重复声明。

(1)人口总数预测。图 2 和表 4 是人口总数预测的结果。参考模型二和准确结果(参考图 5 模型二与模型一 未来 15 年的人口总数预测结果比对图)对数据进行分析我们发现短期内人口的预测总数量与准确结果吻合的比较好。但当年份 t 大于 2013 时，我们发现模型所预测的人口数要大于准确结果。我们认为这是由于模型把男女人口分为城、镇、乡三大类后，并未考虑这三个群体之间的人口流动。当前环境下，中国的城镇化水平是逐渐提高的。而由于生活观念的差异与计划生育的普及情况城镇妇女生育率要低于乡村妇女的生育率。这样在城镇化水平不断提高的环境下(即乡村人口向城镇转移)，实际的全国人口增长率必然要低于不考虑城乡人口转移的模型的预测结果，但这种差异是要在一段时间的积累作用下才会明显的显示出来的。所以在模型一所预测的未来人口数图象中，当 t 大于 2013 时，明显的出现预测结果较高的现象。

(2)人口性别比的预测。从图 3 可见其与准确结果期望不同。预测结果男女的不平衡日趋严重。这是由于对男女比例的产生主要影响的因素是出生性别比(见图 7 未来 15 年的出生性别比预测结果柱形图)，而我们在建立模型时，对出生性别比的处理方法只是按现在的数据进行预测而并未加以限制。所以导致不平衡的现象日趋严重。这也说明

政府想要对性别比过高问题进行抑制，采取相应政策保证出生男女比有所改善是必要的。

(3) 城镇化水平的预测，模型一并未考虑城乡人口的流动转移。所以其城镇化水平的变化并无实际意义，在此不做分析。

(4) 由于模型一很多数据不准确，所以在此不再对其进行分析。

5.2 模型二(城乡转移矩阵模型)

5.2.0 基本符号约定

$s=1$ 表示男性；

$s=2$ 表示女性；

$f=1$ 表示从乡村转移到城市；

$f=2$ 表示从乡村转移到镇。

5.2.1 基本思路

在模型一中我们忽略了人口的城乡流动转移建立了改良的矩阵人口增长模型。但实际上我们知道由于城乡生育水平和人口生育意愿的客观差异，人口的城乡流动转移对生育率无疑将会有影响，同时人口的城乡转移会直接影响城乡人口的年龄结构，而且人口城镇化的速度决定了乡城人口年龄转移率的水平。为了准确预测和判断以上三个方面的影响在我们引入 $W_i(t)$ 人口城乡转移比例矩阵进行估算。我们在设定参数时，曾经考虑添加镇县向城市转移的情况并设为 $f=3$ 的情况。然而我们发现， $f=3$ 的情况事实上是被 $f=1$ 和 $f=2$ 两种情况所包含的，这样多设出 $f=3$ 的情况将会导致多余变量的产生，出现冗余量，甚至在一些方程的求解中会因此而出现解不出答案的情况。事实也证明当我们设定 $f=1, f=2, f=3$ 三种状态时，对于 $a_f^s(k)$ 的求解是无法进行的(变量推导时得到的六个式子由于实际上有式子是等价的所以不足以求解方程组)。所以，我们仅设定村乡向城市转移的情况($f=1$)和村乡向镇县转移的情况($f=2$)。

5.2.2 模型的建立

(1) 转移矩阵的建立。转移矩阵 $W_f^s(t)$ 如下：

$$W_f^s(t) = \begin{bmatrix} w_f^s(0,t) \\ w_f^s(1,t) \\ \vdots \\ w_f^s(90,t) \end{bmatrix}$$

其中， $w_f^s(k,t)$ 表示分性别的 f 方向的人口转移比。

(2) 城镇转移矩阵模型的建立。添加人口城乡转移对人口预测的影响，使用转移矩阵计算具体的人数量比例的变化，得到如下模型

$$\left\{ \begin{array}{l} X_1(t+1) = A_1(t)X_1(t) + \beta_1(t)B_1(t)X_1(t)\frac{p_1(t)}{p_1(t)+1} + W_1^1(t) \bullet X_5(t) \\ X_2(t+1) = A_2(t)X_2(t) + \beta_1(t)B_1(t)X_2(t)\frac{1}{p_1(t)+1} + W_1^2(t) \bullet X_6(t) \\ X_3(t+1) = A_3(t)X_3(t) + \beta_2(t)B_2(t)X_3(t)\frac{p_2(t)}{p_2(t)+1} + W_2^1(t) \bullet X_5(t) \\ X_4(t+1) = A_4(t)X_4(t) + \beta_2(t)B_2(t)X_4(t)\frac{1}{p_2(t)+1} + W_2^2(t) \bullet X_6(t) \\ X_5(t+1) = A_5(t)X_5(t) + \beta_3(t)B_3(t)X_5(t)\frac{p_3(t)}{p_3(t)+1} - W_1^1(t) \bullet X_5(t) - W_2^1(t) \bullet X_5(t) \\ X_6(t+1) = A_6(t)X_6(t) + \beta_3(t)B_3(t)X_6(t)\frac{1}{p_3(t)+1} - W_1^2(t) \bullet X_6(t) - W_2^2(t) \bullet X_6(t) \end{array} \right.$$

式中，转移矩阵 $W_f^s(t)$ 和人口总数向量 $X_i(t)$ 的点乘积代表了对于 t 年 i 类人，人口流动引起的人口转移的数量；其他符号代表含义与模型一相同。

5.2.3 模型的求解

模型二建立在模型一的基础上，所以对模型的求解是在模型一结果以求得的基础上进行的，也就是说我们对模型二的求解就是对其转移矩阵预测的求解，得到转移矩阵即可得到模型二。我们用 $w_{fs}^g(k)$ 表示乡-镇-城转移人口的年龄模式，即 k 岁分性别的 f 方向的人口转移率占该类人口总转移率 $Sw_f^s(t)$ 的比例。从实际数据出发我们可以看到在城镇化的过程中，人口的年龄模式是基本稳定的，所以我们假设 $w_{fs}^g(k)$ 与年龄、类型有关而与具体年份无关，这个假设在在经济环境比较稳定的前提下是合理的。通过数据的统计我们可以得人口总转移率 $Sw_f^s(t)$ ，结合转移人口的年龄模式我们可以得到 $W_f^s(t)$ 的各项 $w_f^s(k, t) = Sw_f^s(t) w_{fs}^g(k)$ ，于是我们求解的转移矩阵的问题转化为对 $Sw_f^s(t)$ 和 $W_f^s(t)$ 的求解。

(1) $w_{fs}^g(k)$ 的预测估计。设 t 年 k 岁 i 类人口转移数为 $a_f^s(k, t)$ ，在无人人口转移条件下，由 t 年人口推算 $t+1$ 年人口应为 $x_{i(k+1)}(t+1) = s_{ik}(t) x_{ik}(t)$ ，从实际出发考虑 $a_{ik}(t)$ 得到

$$\left\{ \begin{array}{l} x_{1(m+1)}(t+1) = s_{1m}(t)x_{1m}(t) + a_1^1(m+1, t+1) \\ x_{2(m+1)}(t+1) = s_{2m}(t)x_{2m}(t) + a_1^2(m+1, t+1) \\ x_{3(m+1)}(t+1) = s_{3m}(t)x_{3m}(t) + a_2^1(m+1, t+1) \\ x_{4(m+1)}(t+1) = s_{4m}(t)x_{4m}(t) + a_2^2(m+1, t+1) \\ x_{5(m+1)}(t+1) = s_{5m}(t)x_{5m}(t) - a_1^1(m+1, t+1) - a_2^1(m+1, t+1) \\ x_{6(m+1)}(t+1) = s_{6m}(t)x_{6m}(t) - a_1^2(m+1, t+1) - a_2^2(m+1, t+1) \end{array} \right.$$

式中， $s_{ik}(t)$ 表示 t 年 k 岁 i 类人的存活率。

6个式子但实际上其直间是存在等价的式子的任取四个式子进行求解。可以得到 t 年时 k (不包括0)岁 i 类人口转移数 $a_f^s(k)$ ，再利用自动平滑法推出 $k=0$ 时的 $a_f^s(0,t)$ 即得到 t 年时全部的 $a_f^s(k)$ 。

$$\text{由于} \begin{cases} w_1^s(k,t) = \frac{a_1^s(k,t)}{x_{(4+s)k}(t) + a_1^s(k,t) + a_2^s(k,t)} \\ w_2^s(k,t) = \frac{a_2^s(k,t)}{x_{(4+s)k}(t) + a_1^s(k,t) + a_2^s(k,t)} \end{cases} \text{通过已知数据我们可以得到}$$

$$t = \{2002, 2003, 2004, 2005\} \text{时的} w_f^s(k,t); \text{且有 } Sw_f^s(t) = \sum_{k=0}^{90} w_f^s(k,t)$$

$$\text{于是可以求出 } w_{fs}^s(k) = \frac{\frac{w_f^s(k,2002)}{Sw_f^s(2002)} + \frac{w_f^s(k,2003)}{Sw_f^s(2003)} + \frac{w_f^s(k,2004)}{Sw_f^s(2004)} + \frac{w_f^s(k,2005)}{Sw_f^s(2005)}}{4}$$

(2)对城镇化水平是预测。城镇化水平的预测可看作是对 $Sw_f^s(t)$ 预测的数据准备。

城镇化水平 $u(t)$ 经过回归分析，不同模型的比对我们发现我使用美国地理学家诺瑟姆(Ray. M. Northam)所概括的 S 型曲线—— $u(t) = \frac{1}{1 + e^{\alpha + \beta t}}$ (α 和 β 为待定参数)对中国的 $u(t)$ 进行预测效果最好，对其进行统计检验得到 $R^2=0.968075326$ ，拟合效果较好。可见我们使用这样的分析是合理的

设城镇化水平在 t 年时为 $u(t)$ ， $u(t)$ 即城镇人口总数和总人数的比，我们可以将其分解为 $u(t) = u_1(t) + u_2(t)$ 其中 $u_1(t)$ 代表城镇化水平即城市人口占总人数的比例， $u_2(t)$ 代表镇化水平即镇人口数占总人数的比例。所以 $u_3(t) = 1 - u_1(t) - u_2(t)$ ，将 $u_1(t)$ 、 $u_2(t)$ 等再次分解得到 $u_{1s}(t)$ 、 $u_{2s}(t)$ 、 $u_{3s}(t)$ 为下一步对 $Sw_f^s(t)$ 预测做准备。

(3)对 $Sw_f^s(t)$ 进行预测计算。结合已经预测的数据建立了对 $Sw_f^s(t)$ 的数据预测模型由已知得，

$$\begin{cases} a_1^s(t) = u_{1s}(t+1)[X(t) + B(t+1) - D(t+1)] - [x_1^s(t) + b_1^s(t+1) - d_1^s(t+1)] \\ a_2^s(t) = u_{2s}(t+1)[X(t) + B(t+1) - D(t+1)] - [x_2^s(t) + b_2^s(t+1) - d_2^s(t+1)] \\ -a_1^s(t) - a_2^s(t) = u_{3s}(t+1)[X(t) + B(t+1) - D(t+1)] - [x_3^s(t) + b_3^s(t+1) - d_3^s(t+1)] \end{cases}$$

式中, $a_f^s(t)$ 表示 t 年分性别 f 方向的人口转移量总数; $X(t)$ 为 t 年总人口; $B(t+1)$ 为 $t+1$ 年出生人口总数; $D(t+1)$ 为 $t+1$ 年死亡人口总数; $x_j^s(t)$ 为 t 年 i 类人的总人口; $b_j^s(t+1)$ 为 $t+1$ 年 i 类人的出生人口总数; $d_j^s(t+1)$ 为 $t+1$ 年 i 类人的死亡人口总数; i 由 s 和 j 确定。可以计算得到 $a_f^s(t)$ 。

$$\text{同时有} \begin{cases} a_{1s}^g(t) = \sum_{k=0}^{90} [w_{1s}^g(k) x_{(4+s)k}^g(t)] \\ a_{2s}^g(t) = \sum_{k=0}^{90} [w_{2s}^g(k) x_{(4+s)k}^g(t)] \end{cases}, \text{式中, } a_{fs}^g(t) \text{ 为 } t \text{ 年假想队列乡-城转移人口的}$$

规模。可以计算得到 $a_{fs}^g(t)$ 。从而我们可以得到 $Sw_f^s(t) = \frac{a_{fs}^g(t)}{a_{fs}^g(t)}$ 。

(4) 由 $w_f^s(k, t) = Sw_f^s(t) w_{fs}^g(k)$ 可得出矩阵 $W_f^s(t)$ 的所有元素, 矩阵求解出

(5) 根据得到的转移矩阵和模型二得到的人口变化矩阵可以得到新的流动的人口增长模型。可以进行相关预测。然后对数据进行统计整理得到以下结果: 图 4 模型二未来 15 年的人口总数预测结果柱形图, 图 5 模型二与模型一未来 15 年的人口总数预测结果比对图, 图 6 模型二未来 15 年的人口性别比预测结果柱形图, 图 7 未来 15 年的出生性别比预测结果柱形图, 图 8 模型二未来 15 年的人口城镇化水平, 图 9 模型二未来 15 年的老龄化预测, 图 10 模型二未来 15 年的生育旺盛期妇女人数的预测, 图 11 模型二未来 15 年的育龄妇女人数的预测

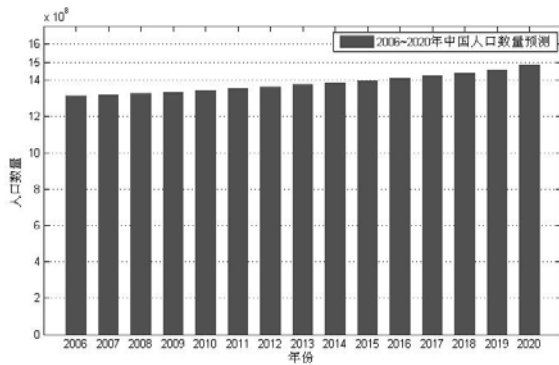


图 4 模型二 未来 15 年的人口总数预测结果柱形图

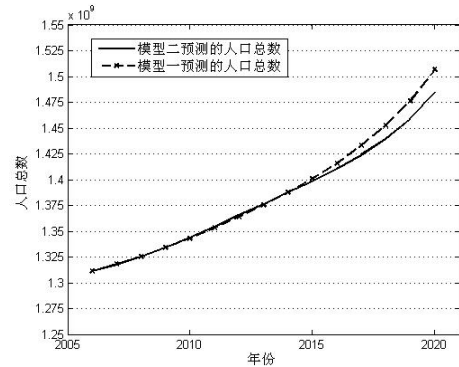


图 5 模型二与模型一 未来 15 年的人口总数预测结果比对图

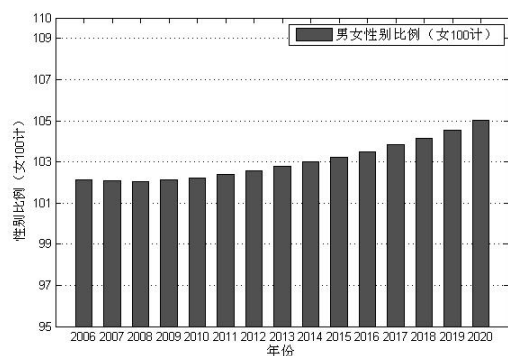


图 6 模型二 未来 15 年的人口性别比预测结果柱形图

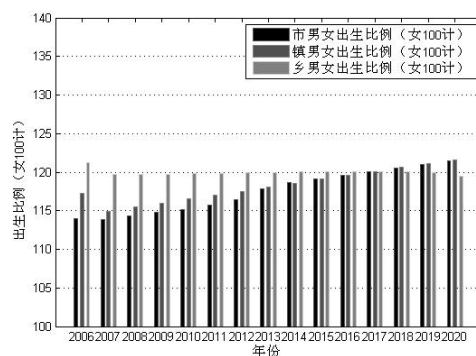


图 7 未来 15 年的出生性别比预测结果柱形图

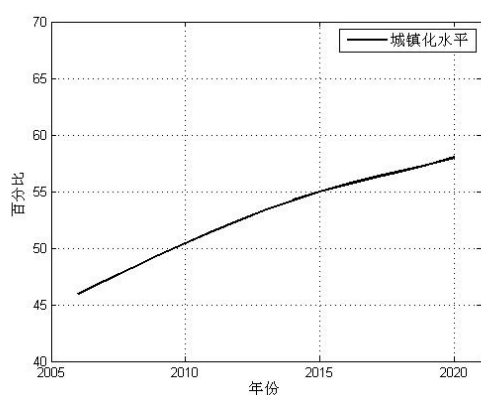


图 8 模型二 未来 15 年的人口城镇化水平预测

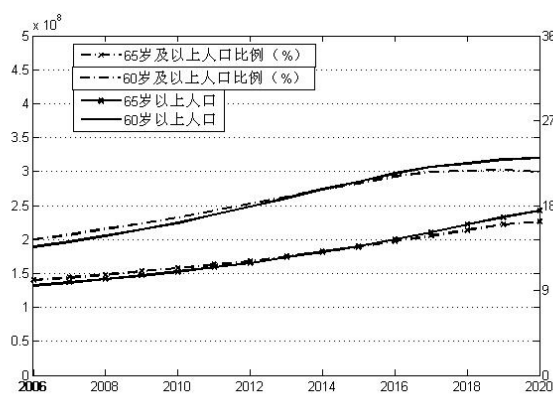


图 9 模型二 未来 15 年的老龄化预测

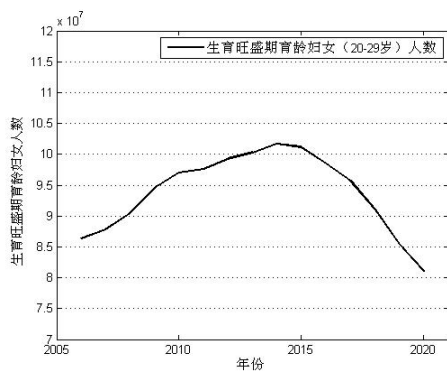


图 10 模型二 未来 15 年的生育旺盛期妇女人数的预测

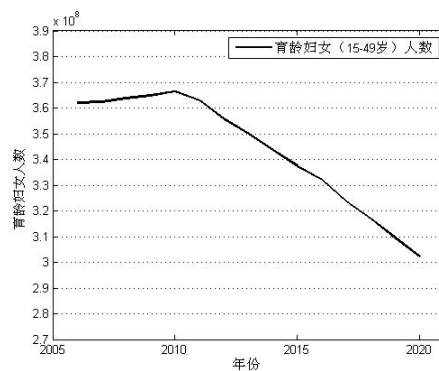


图 11 模型二 未来 15 年的育龄妇女人数的预测

5.2.4 结果分析

(1) 人口总数预测。图 4 和表 1 表现了我们得到了 15 年内中国人口总数预测值。可以看出其与准确结果吻合的较好。且对图 5 的分析可知由于引入了城乡人口移动的影响，模型二的结果不会出现模型一中的后期数量过大的问题，也进一步说明了模型一的不足之处和模型二的优越之处

(2) 人口性别比和出生性别比的预测。从图 6 和图 7 可见其与准确结果期望不同。预测结果男女的不平衡日趋严重。模型一和模型二对出生性别比的预测是使用同一模型

的(使用了 BP 神经网络对其进行的预测)。同时流动的男女比例与原群的性别比大体一致,即引入转移矩阵对性别比的影响不大。所以模型一和模型二所得结果基本相同。

(3) 城镇化水平的预测。图 8 为预测结果,其结果与准确结果进行比对较为准确,虽然有一定差别,但差别处于可以接受的范围内。准确的城镇化水平预测,说明在模型中引入转移矩阵对人口城镇转移的处理是科学合理的。

(4) 老龄化程度的预测。图 9 作为预测结果与准确结果进行比对可见其准确性是比较好的。

(5) 育龄妇女人数和生育旺盛期妇女的人数的预测。图 10 和图 11 通过比对可以得出两者的预测结果都是比较好的。于是结合老龄化的预测我们知道模型二对于分年龄的预测是比较准确的。

(6) 综上所述,模型二的数据预测在中短期内更有指导意义,因为;①由模型一发展而来 ②引入了新的因素转移矩阵适应实际情况 ③更准确,更可靠,更全面。所以模型二是一个很好的中短期人口增长预测模型

5.3 模型三(长期人口增长预测模型)

5.3.1 基本思路

由于以上模型由于其数据预测方法上的局限使得在 50 年以上的人口长期增长的预测并不够准确。于是我们建立了长期人口增长预测模型使之能够更好的进行长期预测。

根据模糊数学的理论我们知道,对于一个复杂模型过分要求精确反而模糊,而适当的模糊却可以达到精确的结果。对于对人口进行长期预测这个复杂的问题对初始数据的要求并不是越精越好而是应对原始数据进行一系列的模糊处理。通过对中国人口政策和所有的人口统计数据我们可以知道在长期预测中女性的数据变化的可信度要比男性的高所以我们采用 Leslie 提出描述女性人口变化规律的模型为基础采取对女性变化进行预测从而得到整体变化规律的处理方法。以下所给的数据都是针对女性的。

5.3.2 模型的建立

(1) 数据处理。我们知道按照目前人的价值观来看乡村人口转移到城镇是总体趋势,且在经济发展水平较为稳定的情况下,该趋势将会保持比较长的时间。所以我们从乡村人口城镇化角度出发将原有的分为城、镇、乡三类的人口重新划为乡村和城镇两类数据,以便推算出长期的城镇化的做法是合理的。

(2) 死亡率的预测。我们考虑选用灰色预测模型或随机预测模型对死亡率进行预测,由于要预测长期的人口增长,而灰色预测模型相对于随机预测模型来说不适宜用于长期数据的预测。所以针对长期模型我们采用 Lee-Carter 预测模型对人口的死亡率进行预测。模型如下:

$$\ln[m_x(t)] = a_x + b_x k(t) + \varepsilon_x(t)$$

式中, $m_x(t)$ 是 t 年 x 岁人的死亡率; $k(t)$ 是时间因子,作为测量死亡率整体水平的指标; a_x 是年龄别的死亡率取自然对数以后的平均值,描述了死亡率年龄模式的时间趋势; b_x 是年龄因子,用来描述当时间因子变化时,各个年龄组死亡率相对变化的系数; $\varepsilon_x(t)$ 表示。 t 年 x 岁死亡率的残差。

(3) 生育率的预测。考虑我国计划生育政策的全面实施,以及人们价值观念的变化

趋势我们可以推断中国未来城市和农村的生育率水平是保持稳定的。于是我们得到方程

$$b_r(k, t) = R_r(k)(1 - u(t)) + U_r(k)u(t) \quad k = \{0, 1, 2, \dots, 90\}$$

式中 $b_r(k, t)$ 表示 t 年, 年龄为 k 的人的生育率; $u(t)$ 表示 t 年的城乡人口比 (即城镇化水平); $R_r(k)$ 表示乡村年龄为 k 的人的生育率; $U_r(k)$ 表示城市年龄为 k 的人的生育率。

(4) 城镇化水平的预测。我们根据美国地理学家诺瑟姆 (Ray. M. Northam) 所概括的 S 型曲线—— $u(t) = \frac{1}{1 + e^{\alpha + \beta t}}$ (α 和 β 为待定参数) 对中国的 $u(t)$ 进行预测。

(5) 女性人口增长长期预测模型。经过上述数据的预测我们可以得到长期的预测模型如下

$$\begin{cases} p(0, t) = \sum_{k=15}^{49} [b_r(k, t)p(k, t)] \\ p(k, t) = [1 - d_r(k-1, t)]p(k-1, t) \end{cases}$$

到此我们得到了完整的女性人口增长长期预测模型。

(6) 男女性别比的预测。在完成女性模型的基础上我们要对中国总人口的男女性别比 $P_{mw}(t)$ 进行预测从而才能得到全国总人数的预测。

设 t 年男女的生育比为 $K(t)$, 女性人口数为 $p_w(t)$, 男性人口数为 $p_m(t)$, 女性的生产率为 $g(t)$ 。于是可以得到

$$\begin{cases} \frac{dp_m(t)}{dt} = p_w(t)g(t)\frac{K(t)}{1 + K(t)} \\ \frac{dp_w(t)}{dt} = p_w(t)g(t)\frac{1}{1 + K(t)} \end{cases}$$

于是由 $P_{mw}(t) = \frac{p_m(t)}{p_w(t)}$ 及上面的数据可以得到 $P_{mw}(t)$ 的预测曲线。

5.3.3 模型求解

(1) 数据处理

(2) 死亡率的预测。 a_x 可由已知数据的统计得到; 通过强制 b_x 总和为 1, $k(t)$ 总和为 0 的限制条件, 采用奇异值分解 (SVD) 方法求解得到 b_x 和 $k(t)$ 。

(3) 生育率的预测。 $R_r(k)$ 、 $U_r(k)$ 的值由于是逐年最终趋于平衡的值我们取已知数据的平均值为该值。

(4) 城镇化水平的预测。对方程变形可以得到 $\ln\left(\frac{1}{y} - 1\right) = \alpha + \beta t$ 使用普通最小二

乘法(OLS)进行回归分析得到 $\alpha = 0.1645, \beta = -0.03457$ ，对得到的结果进行统计分析得到 $R^2=0.968075326$ ，说明我们用这条曲线是合理的准确的。

(5) 女性人口增长长期预测模型。对差分方程进行递推求解。

(6) 男女性别比的预测。

(7) 求解出预测结果为图 12 未来 50 年人口总数和劳动年龄人口预测，图 13 未来 50 年老龄化程度的预测，图 14 未来 50 年生育旺盛期妇女人口的预测，图 15 未来 50 年育龄妇女人数的预测，表 5 未来 50 年人口总数预测结果。

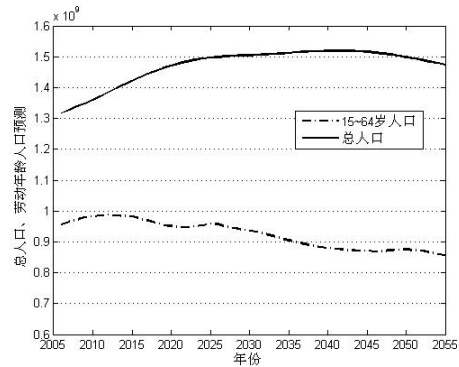


图 12 未来 50 年人口总数和劳动年龄人口预测

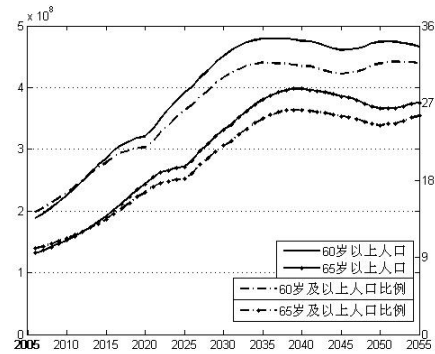


图 13 未来 50 年老龄化程度的预测

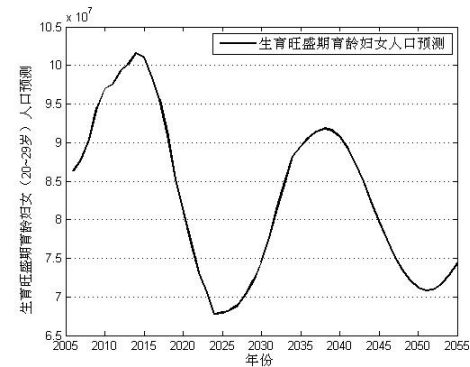


图 14 未来 50 年生育旺盛期妇女人口的预测

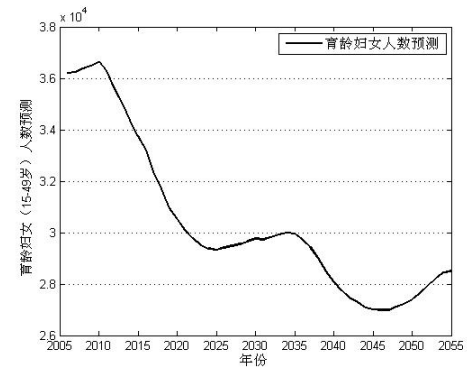


图 15 未来 50 年育龄妇女人数的预测

表 5 未来 50 年人口总数预测结果

单位：十亿人

年份	人口数量	年份	人口数量	年份	人口数量	年份	人口数量
2006	1.3178	2007	1.3280	2008	1.3385	2009	1.3493
2010	1.3606	2011	1.3726	2012	1.3850	2013	1.3977
2014	1.4102	2015	1.4223	2016	1.4337	2017	1.4444
2018	1.4542	2019	1.4632	2020	1.4712	2021	1.4783
2022	1.4844	2023	1.4895	2024	1.4935	2025	1.4967
2026	1.4990	2027	1.5009	2028	1.5023	2029	1.5037

2030	1.5049	2031	1.5062	2032	1.5075	2033	1.5088
2034	1.5104	2035	1.5121	2036	1.5138	2037	1.5154
2038	1.5168	2039	1.5180	2040	1.5189	2041	1.5194
2042	1.5194	2043	1.5189	2044	1.5179	2045	1.5161
2046	1.5136	2047	1.5107	2048	1.5071	2049	1.5032
2050	1.4988	2051	1.4941	2052	1.4892	2053	1.4841
2054	1.4788	2055	1.4737				

5.3.4 结果分析

(1) 长期人口总数和劳动年龄人口的预测。图 12 与准确结果进行比对可以看到预测的结果是很准确的。人口总数持续增长趋于稳定，并在 2045 年后出现了回落。

(2) 长期老龄化程度的预测。图 13 与准确结果基本相同，预测准确。老龄化速度在加快，老龄化程度日见严重。至到 2030 年后逐渐稳定。

(3) 长期育龄妇女人数和生育旺盛期妇女的人数的预测。图 14 和图 15 通过比对可以得出两者的预测结果都是比较好的。得出年龄的预测是比较准确的结论。当然其数据本身对人口数量的变化也有着很重要的影响。

(4) 综上所述，长期模型的预测结果贴近实际要求、准确科学、可信度高是一个优良的模型。

5.4 模型四(综合评价模型)

5.4.1 基本思路

基于层次分析法我们建立一个对于各个人口增长模型进行综合评价的模型，此模型的层次结构图下图所示：

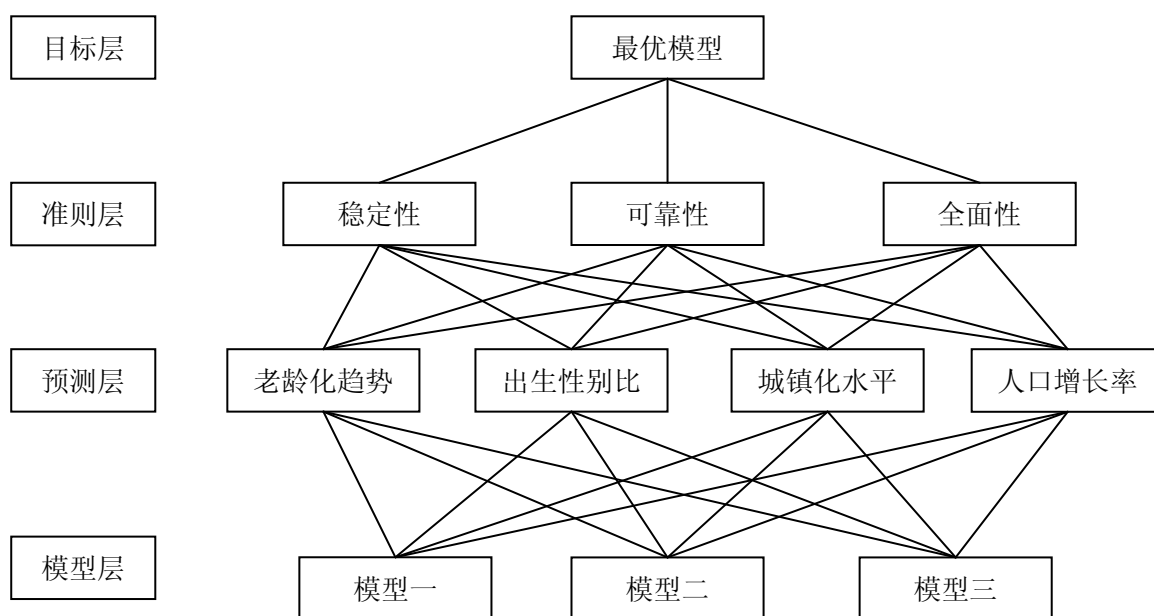


图 16 综合评价层次结构图

稳定性：指在预测过程中原始数据或中间数据的误差和实际过程中偶然事件的发生

对最终结果的影响程度；

可靠性：预测结果与实际情况的吻合程度，即预测的准确程度；

全面性：模型在充分的利用以有数据对问题进行多方面的综合建模的程度。

5.4.2 权重向量的确定

(1) 准则层对目标层权向量的确定。由准则层对目标层的影响，我们知道模型的稳定性和可靠性相对来说比全面性要重要。不妨设成对比矩阵为

$$O = \begin{bmatrix} 1 & 1 & 2 \\ 1 & 1 & 2 \\ \frac{1}{2} & \frac{1}{2} & 1 \end{bmatrix}, \text{ 求得权重向量 } w^{(2)} = (0.4, 0.4, 0.2)^T$$

(2) 预测层对准则层权向量的确定。同理由预测层的影响我们发现各个预测结果对准则层来说是几乎同等重要的，于是成对比矩阵为

$$A_1 = A_2 = A_3 = \begin{bmatrix} 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 \end{bmatrix}, \text{ 求得所有权向量均为 } w_k^{(3)} = (0.5, 0.25, 0.25, 0.25)^T \quad (k=1,2,3)$$

(3) 模型层对预测层权向量的确定。在求模型层对于预测层的权向量时，我们知道预测层对模型层的评价是具有模糊特征，由于①因为评价因素的多元性，即确定模型的预测层权重要考虑老龄化趋势、出生性别比、城镇化水平、人口增长率等诸多因素。②评价因素的模糊性，有些指标的数字化并不是确定的，很难量化估计。③评价标准的多样性，对于同一个评价因素对于不同的人不同的目标其评判标准可能是不同的。基于以上评价对象的基本特征，我们认为应该用模糊评价的方法对其出比较客观、准确和公正的评价。

通过查找资料确定各种科学的，不同评判标准进行我们综合标准可以进行组别化，按下表进行评价

表 6 一评判标准对三个模型的分组评判

指标	老龄化趋势					出生性别比					城镇化水平					人口增长率				
分组	优	良	中	及	差	优	良	中	及	差	优	良	中	及	差	优	良	中	及	差
模型一																				
模型二																				
模型三																				

对于某模型某方面被化为优、良、中、及、差的个数计做

$$X_{ij} = [x_{ij}(1) \quad x_{ij}(2) \quad x_{ij}(3) \quad x_{ij}(4) \quad x_{ij}(5)]$$

其中， $x_{ij}(k)$ 表示第*i*个模型的第*j*个方面中第*k*项的个数。令 $w = (0.25, 0.5, 0.75, 1.00)^T$ 和

$$X_{ij} = x_{ij} w \text{ 即可得到 } w_k^{(4)} = \left(\frac{x_{1k}}{\sum_{i=1}^3 x_{ik}}, \frac{x_{2k}}{\sum_{i=1}^3 x_{ik}}, \frac{x_{3k}}{\sum_{i=1}^3 x_{ik}} \right)。$$

5.4.3 整体模型的建立求解

最后令 $W^{(i)} = [w_1^{(i)} \quad w_2^{(i)} \quad \dots \quad w_n^{(i)}]$ 其中 n 为上一层次元素个数。这样我们通过 $w^{(s)} = W^{(s)}W^{(s-1)} \dots W^{(3)}w^{(2)}$ 即可判断出最优模型, 由于对模型层求权向量的时候使用的是模糊评价的方法。

在对其它层求权向量时采用的是一致阵作为成对比较矩阵。所以各个权向量及组合权向量可以通过一致性检验和组合一致性检验。

5.4.4 结果分析

由于时间原因我们并未能进行的具体的模型求解。但从理论上来说该模型是可以得到较准确的评价结论的。

6、模型评价

6.1 模型一

该模型综合考虑了出生性别比例以及年龄组成比例等因素进行离散形式的分组人口预测。具有对男性数据同步预测结果较为准确、充分利用数据、离散形式进行人口预测便于计算机作数值计算、数据预测结果贴近现实等优点, 但模型由于并未考虑人口流动因素对人口增长的影响即无法考虑到城市化水平这一重要因素对人口的影响, 所以只能作为比较优化的粗略中短期模型。

通过模型求解, 我们发现该模型的数据预测结果是较为贴近现实的。其预测结果可以经过整理进而得到包括人口总数、性别比例、老龄化程度等多个重要的预测指标能对人口增长进行比较完全的预测。

然而要想更加精确的对人口中短期进行预测就必须考虑城、镇、乡三个地域的人口流动必须考虑城市化对人口的影响, 这就要求了一个更佳的模型进行预测。于是我们在原模型的基础上引入了人口转移矩阵, 建立了模型二。

6.2 模型二

该模型是在模型一的基础上生成的, 它包含了模型一的全部优点。能对人口总数、性别比例、老龄化程度等多个重要的数值进行预测。同时由于人口转移矩阵的引入, 使得它涵盖了城、镇、乡三地人口流动信息, 包含了城市化水平的重要因素, 因而其有能力对未来中短期人口进行更全面的预测。

但是事实上在长期的人口增长中, 人们的生活环境会发生一定的变化, 影响生育率和死亡率的因素都在长时间的过程中产生或多或少的变化。且随机的事件都会对未来人口造成一定的影响, 这就使得无论是模型一还是模型二都无法很准确地预测长期的人口变化。

这样, 我们就必须从另外一个角度建立一个预测长期人口变化的模型。

6.3 模型三

由于我们在建立模型一和模型二过程中用神经网络和灰色预测的方法预测了死亡率等一些重要数据, 所以由此数据而得到的预测数据不能保证绝对的准确。并且由于模型一和模型二是以年为单位的依次叠代预测, 使得在长期过程中, 分别预测男性与女性的人口波动性将大于仅预测单一性别人口的波动性。这样, 我们不妨只是建立一个预测女性的模型, 再建立模型对未来男女比例进行预测, 这样就可得到全部的人口变化。长远来看, 这样求得的预测结果要比分别求男性女性而得到的预测结果误差要小。

由于用差分方程递推求解有着时间越靠后准确性越低的缺点, 为了克服这一缺点,

我们对部分数据进行了模糊处理，使之更适用于长期的预测分析。（由于微分方程建立的人口模型对于年龄结构的描述不如差分方程建立的模型，所以我们仍采用差分方程来描述当前在面临人口老龄化问题的情况下的人口情形）

通过对已有材料的查询以及相关书籍的阅读，我们分别引入了死亡率的经济学模型、生育率的变化模型以及预测城镇化水平的诺瑟姆 S 曲线模型。进行求解后我们发现其预测准确度是很高的。但由于在对男女比例预测模型的建立过程中，我们仅是对进行了相关的简化处理后列出的化简后的微分方程模型进行粗略预测，导致了一定的误差。

总体来说，模型三的建立和求解，很好的帮助我们进行长期的人口变化预测。

6.4 模型四

基于层次分析法和模糊理论思想的对人口模型进行综合评价的模型。它的建立大体上是基于层次分析法的。但是在模型层的权重求解上我们利用了模糊数学的原理求得了该层的权重矩阵。该模型的实用意义在于能够利用层次分析法较为全面的评价各个入口模型。只是在预测层的权重矩阵的建立上以及准则层的权重矩阵的建立上我们用了简化的方式进行了求解，这在一定程度上妨碍了评价的绝对公平性。

7、模型扩展

7.1 基本思路

对于人口增长预测模型从根本上讲，是一个这样的模型：随时间变化的量 X (人口总量) 有输入(出生)和输出(死亡)，有类别(性别、城乡)区别，并且不同类别之间可以相互转化(城乡转移)，模型在已知某点时间前的一系列 X 数据情况下，可以对该时间点后的 X 进行预测，同时也可以对 X 的一些相关数据(性别比、城镇化、老龄化等)进行预测。那么对于任意的 X 的求解都可以套入人口的数据，从而得到人口预测结果。

7.2 模型的建立思想

对于某工厂的所有产品上架数 S 进行预测。 S 代表该工厂所有处于销售中的产品总数相当于人口模型的人口总数，那么工厂生产产品并使其处于销售状态的比率相当于我们人口模型中的生育率，某产品由于质量问题被召回或者不受消费者欢迎赢利过小导致下架的比率相当于人口模型中的死亡率。工厂产品有多种类型不同类型可以对应模型中的分性别城乡人口。同时其产品可能存在着可升级的空间那么产品的升级所带来的变化就相当于城镇人口流动。不同种类的产品的上架数量的比例就是人口模型中的性别比、老龄化程度等指标。

7.3 模型的求解和评价

产品上架数问题在经济领域中对上架问题进行建模，对其进行求解把人口的数据带入即可，由于时间原因我们并未能将模型建立出来并求解。但从基本思路出发将人口模型基本化那么其求解模型是多种多样的这里也只是以上架模型为例。

参考文献

- [1] 姜启源等，数学模型（第三版），北京：高等教育出版社，2003
- [2] 张圣勤，《MATLAB 7.0 实用教程》，北京：机械工业出版社，2006
- [3] 张佰瑞，《城市化水平预测模型的比较研究——对我国 2020 年城市管理研究所》，《理论界》，2007 年 4 月刊：048-051，2007。
- [4] 国家统计局年鉴编辑部，《中国统计年鉴 2006》，北京：国家统计局，2006

附录

表 7 1987-2005 人口数及结构 [4]

本表各年人口未包括香港、澳门特别行政区和台湾省的人口数据。

单位：万人

年 份	年底人口数	按 性 别 分				按 城 乡 分			
		男		女		城 镇		乡 村	
		人口数	比重 (%)	人口数	比重 (%)	人口数	比重 (%)	人口数	比重 (%)
1978	96259	49567	51.49	46692	48.51	17245	17.92	79014	82.08
1980	98705	50785	51.45	47920	48.55	19140	19.39	79565	80.61
1985	105851	54725	51.70	51126	48.30	25094	23.71	80757	76.29
1989	112704	58099	51.55	54605	48.45	29540	26.21	83164	73.79
1990	114333	58904	51.52	55429	48.48	30195	26.41	84138	73.59
1991	115823	59466	51.34	56357	48.66	31203	26.94	84620	73.06
1992	117171	59811	51.05	57360	48.95	32175	27.46	84996	72.54
1993	118517	60472	51.02	58045	48.98	33173	27.99	85344	72.01
1994	119850	61246	51.10	58604	48.90	34169	28.51	85681	71.49
1995	12112	61808	51.03	59313	48.97	35174	29.04	85947	70.96
1996	122389	62200	50.82	60189	49.18	37304	30.48	85085	69.52
1997	123626	63131	51.07	60495	48.93	39449	31.91	84177	68.09
1998	124761	63940	51.25	60821	48.75	41608	33.35	83153	66.65
1999	125786	64692	51.43	61094	48.57	43748	34.78	82038	65.22
2000	126743	65437	51.63	61306	48.37	45906	36.22	80837	63.78
2001	127627	65672	51.46	61955	48.54	48064	37.66	79563	62.34
2002	128453	66115	51.47	62338	48.53	50212	39.09	78241	60.91
2003	129227	66556	51.50	62671	48.50	52376	40.53	76851	59.47
2004	129988	66976	51.52	63012	48.48	54283	41.76	75705	58.24
2005	130756	67375	51.53	63381	48.47	56212	42.99	74544	57.01

注：1. 1982 年以前数据为户籍统计数；1982—1989 年数据根据 1990 年人口普查数据有所调整；1990—2000 年数据根据 2000 年人口普查数据进行了调整；2001—2004 年数据为人口变动情况抽样调查推算数；2005 年数据根据全国 1%人口抽样调查数据推算（下表同）。

2. 年底人口数和按性别分人口中包括中国人民解放军现役军人，按城乡分人口中现役军人计入城镇人口。